Cardiomyocytes from human induced pluripotent stem cells: capturing disease severity of LQT2 syndrome and the impact of chromosome aberrations

Brandao, K.O.

1

General Introduction and thesis outline

Cardiovascular disease (CVD) accounted for 1.8 million deaths across Europe in 2017, with an estimated economic impact of 210 billion euros[1]**.** Among these, cardiac arrhythmias are an important contributor. Arrhythmias, or abnormal heart beats, can be caused by mutations in ion channel genes that regulate electrical behaviour of the heart and subsequently contraction of cardiomyocytes, or by drugs that adversely affect ion channel behaviour. One underlying reason for the failure to find effective drug treatments for such arrhythmias is that the pathophysiology of many cardiac diseases and drug effects on the healthy heart are not well captured in rodents. In the case of ion channel diseases, not only does ion channel expression and activity differ between humans and rodents, but genetic differences mean that variants and mutations causing or predisposing humans to CVD may have little or no effect in transgenic mice even if there is a genetic equivalent. Aside from differences in ion channel physiology, the immune system, inflammation, and genetic diversity all impact severity of the disease phenotype and the effectiveness of new drugs in clinical trials. There is clearly an urgent need for better models that reflect ion channel conditions in humans.

In 2007 we had the first reports describing induced pluripotent stem cells derived from human tissue (hiPSCs)[2], and in a short time, this has become a remarkably powerful technique that allows the reprogramming of somatic cells from any (healthy or disease-bearing) individual to a pluripotent state[3]. Requiring the overexpression of only three transcription factors (SOX2, KLF4, OCT3/4), but benefitting from expression of a fourth (c-myc)[2,4], and with a multitude of reprogramming methods and appropriate primary cell sources available[5,6], hiPSC technology has revolutionised genetic research in general and that of the heart in particular. Using hiPSCs derived from specific patient genetic backgrounds, it is possible to create *in vitro* disease models that reflect the clinical features from a

wide repertoire of monogenic and multifactorial pathologies. These models ultimately contribute to unravelling known and novel disease mechanisms and indicate possible therapeutic approaches, as well as creating opportunities to develop stem cell-based therapies that may eventually be autologous[7].

In the cardiac field, hiPSC technology has already been shown to have a substantial impact on understanding many cardiovascular disorders that have been modelled to date[8]. It is increasingly regarded as a game-changer in providing insight into how the genetic contribution to many diseases is regulated even when this is complex, and it has gone some way into indicating why disease phenotypes are variable even among patients with the same ion channel mutation[9]. The complexity arises in part because there are many disease-causing variants in cardiac-related genes and, in addition, there is often variable expressivity and incomplete penetrance of these genes. Consequently, the functional interpretation of what these variants mean is still rather inconclusive and frequently results in setbacks for diagnosis, risk stratification and treatment in clinical management. This is because patients are classified as suffering from the same disease when the underlying molecular mechanisms can actually be different. Thus, the ability to generate hiPSCs from patients, combined with recent advances in genome editing technologies[10] have made it possible to create hiPSC-derived cardiomyocytes (hiPSC-CMs) from patient lines to investigate the pathogenicity of those variants in a *controlled* genetic background, helping to better understand the onset, progression, and molecular mechanisms of mono- and polygenic disease[11]. Another remarkable application of hiPSC is in drug development and drug screening for toxicity on the heart. In a world where CVDs are a major cause of morbidity and treatments are still palliative and not curative, hiPSC-CMs have become a promising alternative in safety-pharmacology; as a possible useful tool for

proarrhythmic risk prediction in the non-clinical drug development phase, and in pre-clinical drug discovery[12].

INSIGHTS FROM HIPSC-CM MODELS

Just a few years after the first hiPSC generation, the first models of cardiac diseases using hiPSC-CMs derived from patients were established. In 2010, Carvajal-Vergara and colleagues were pioneers in demonstrating that hiPSC-CMs could recapitulate the hypertrophic phenotype and altered RAS/MAPK signalling seen in LEOPARD syndrome. The disease phenotype of this autosomal-dominant disease that causes hypertrophic cardiomyopathy was evident as an enlarged surface area occupied by the LEOPARD hiPSC-CMs on the culture surface compared with an unrelated but healthy control[13]. Whilst the criteria for defining "hypertrophic CMs" in culture are more stringent nowadays[14] and we require different controls (preferably isogenic matched through repair of the mutation), this paper led the way to many subsequent studies. In the same year, the channelopathy LQT1 (long QT syndrome type 1) was modelled and demonstrated that besides being able to recapitulate the electrophysiological characteristics of a disease, hiPSC-CMs could also be used for identifying potentially therapeutic approaches. The hiPSC-CMs derived from two patients with increased susceptibility to catecholamine-induced tachyarrhythmia were shown to be therapeutically modulated by the beta-blocker propranolol[15]. These models were important to validate the use of hiPSC-CMs in studying the effects of diverse human genetic disorders, and within a very short time, many cardiomyopathies, cardiometabolic and arrhythmic disorders were modelled[8].

Although a promising system in the cardiac field, and generally acceptable for many ion channel diseases since an almost complete complement of ion channel genes were expressed quite early during differentiation, researchers were confronted with

the need to overcome the general lack of maturation in hiPSC-CM. These cardiomyocytes are phenotypically similar to early foetal stages of cardiac development (around week 16 of gestation)[16], and this proved to be a hurdle for modelling some diseases accurately; particularly those in which phenotypes in patients only developed postnatally or in adulthood. Many studies tried to mimic the environment of heart development in order to reveal the adult disease phenotype[17–19]**.** However, better methods were still needed to reach the same level of maturation as in neonates[20]**,** and even now adult cardiomyocyte phenotypes are not observed.

LONG QT SYNDROME (LQTS)

Congenital long QT syndromes (LQTS) are life-threatening cardiac diseases characterised by a prolongation of QT interval in an electrocardiogram (ECG). This delayed repolarisation of the ventricular myocardium greatly increases the chance of an abnormal heart rhythm leading to syncope episodes, polymorphic ventricular tachycardia, and sudden cardiac death (Figure 1A)[21]**.** At a cellular level this prolongation is translated as a delay in action potential (AP) repolarisation of cardiomyocytes (Figure 1B)**.** LQTS is one of the most common genetic channelopathies of the heart, with an estimated prevalence of 1:2500 individuals[22].

It is mainly caused by single autosomal-dominant mutations in genes encoding for cardiac ion channels or accessory ion channel subunits. *KCNQ1* (LQT1), *KCNH2* (LQT2) and *SCN5A* (LQT3) are the three major genes responsible for LQTS, representing ~90% of all genotype-positive cases[23]. These subtypes have many clinically relevant genetic variants displaying a broad spectrum of pathogenicity. This clinical heterogeneity ranges from being lifelong asymptomatic, to experiencing life-threatening arrhythmias from birth (Figure 2). It is a serious challenge to diagnose

the condition and carry out risk stratification. Despite the diversity of disease, the number of different treatments available are extremely limited[24].

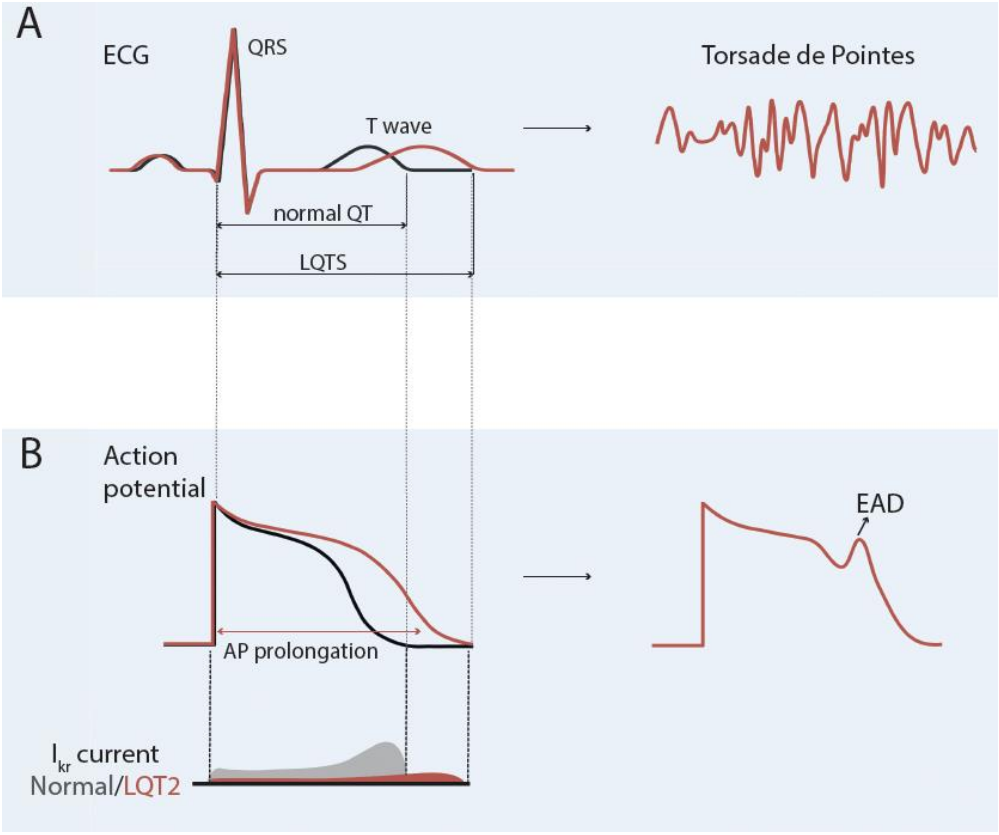

Figure 1: Diagram illustrating the relationship between ventricular action potentials and the electrocardiogram. **(A)** Schematic ECG traces from a healthy (black) and diseased (red) individual. QRS complex and T waves denote the depolarization and repolarization, respectively, of the ventricles. The QT interval is determined by the time of the ventricular repolarisation; its constant prolongation is a risk factor for Torsades de Pointes, shown on the right. The development of the QRS complex and the prolongation of the QT interval are the result of the rapid upstroke of the ventricular action potentials and an increase in the duration of the AP, respectively. **(B)** Schematic traces of normal and prolonged ventricular action potential in LQT2. The reduction of I_{Kr}, caused by mutations in *KCNH2*, delays the repolarization phase of the AP, increasing the time to re-establish the membrane resting potential. Shown on the right is an EAD occurring during the repolarization phase due to AP prolongation.

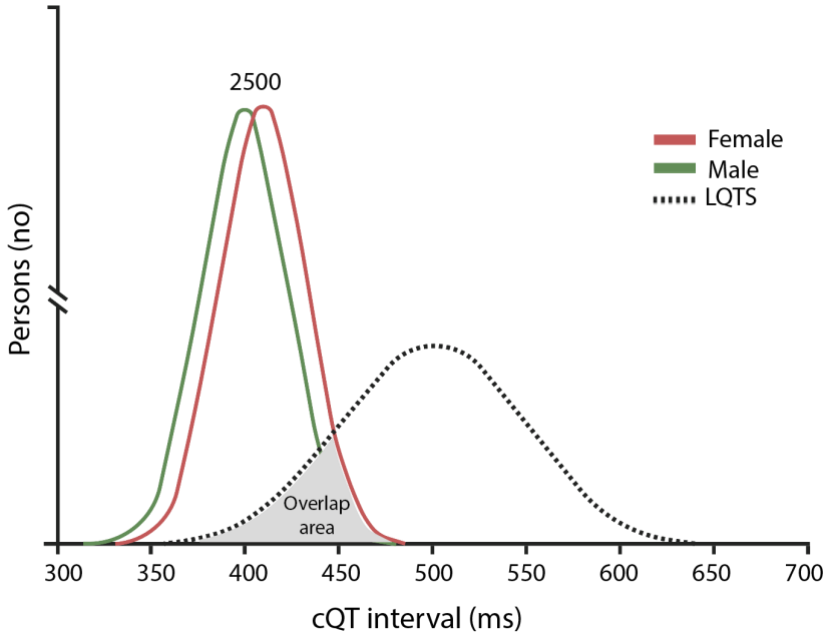

Figure 2: Distribution of QTc values among healthy and disease that shows a subset of patients having concealed LQTS (grey area). Note the break in the y-axis, which reflects the fact that the LQTS vs normal distribution curves must consider the 1-in-2000 estimated incidence of LQTS. Adapted from Taggart et al.[39]**.**

hERG CHANNEL AND LQT2

Much of the focus of this thesis is on Long QT type 2 (LQT2). LQT2 accounts for ~25-40% of all LQTS and most cardiac events in these patients occur during periods of emotional stress or in response to auditory stimuli, often in combination with arousal from sleep. LQT2 is caused by heterozygous loss-of-function mutations in the *KCNH2* gene, also known as *hERG* (human ether-a-go-go-related gene)[24]**.** *KCNH2* encodes the alpha subunit of potassium channel Kv11.1 (hERG), which is responsible for conducting I_{Kr}, the rapid delayed rectifier potassium current[26]. I_{Kr} is a repolarisation current, and its reduction contributes to prolongation of AP duration

(APD) which can trigger early afterdepolarizations (EAD) (Figure 1B). To form the functional potassium channel, KCNH2 subunits oligomerize to form a tetramer. Each subunit contains large cytoplasmic NH2-terminal and COOH-terminal domains and 6 α-helical transmembrane segments (S1-S6). S1-S4 contributes to the voltage sensor domain and S5-S6 in combination with the extracellular pore loop assist to form the K^+-selective pore domain[27] (Figure 3A).

As seen with other LQT types, clinical management of LQT2 can be challenging due to the variable expressivity and incomplete penetrance even among carriers with the same mutation. Genetic factors play important roles in the variability in clinical manifestation[24]**.** To date, more than 300 pathogenic variants in *KCNH2* have been described on ClinVar (public database reporting human genotype-phenotype relationship) plus 794 variants of uncertain significance (VUS). Furthermore, several studies have demonstrated that the location of the primary mutation within the ion channel can influence arrhythmic risk. Usually, patients harbouring mutations in the pore-loop region of the protein are at markedly increased risk for arrhythmia-related cardiac events compared with patients with non-pore mutations[28,29](Figure 3B). In addition, the presence of other intra- or intergenic modifiers can also contribute to phenotypic variation[30]. Thus, development of *in vitro* models that reflect all genetic components present in the disorder could help to evaluate the severity of individual modifications and that would not only have significant diagnostic implications, but also prognostic and therapeutic value.

Many hiPSC lines have been derived from LQT2 patients with mutations in various regions of KCNH2[8]. While these models were important to show the feasibility of modelling LQT2 using hiPSC-CMs, for many comparisons, the lines (control and disease) were from different individuals, and thus they harbour additional genetic variants that could influence the overall disease phenotype. Moreover, huge

phenotypic variability has been observed among hiPSC-CMs from healthy control lines[31].

With the advent of gene editing technologies, it has become possible to avoid this interline variability by using as the control the genetically corrected line from the patient. This in principle eliminates the confounding factors arising from hiPSC derivation and genetic background and allows the effect of a specific variant to be determined. However, with hundreds of *KCNH2* mutations plus VUS and gene modifiers[30,32], the field is in need of a powerful and efficient platform, probably combining in vitro and in silico models, to rapidly assess the phenotype of a large multiplicity of mutations.

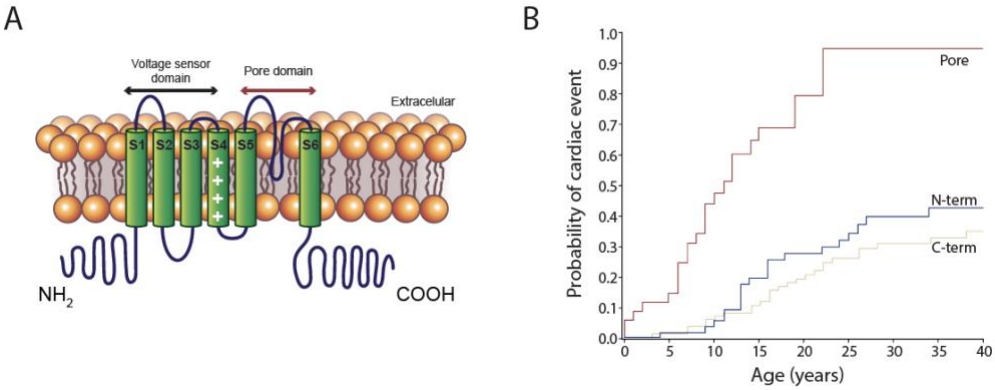

Figure 3: Genotype-phenotype aspect of LQT2. **(A)** Structure of a single alpha subunit of the potassium ion channel KCNH2 indicating the six transmembrane segments (S1–S6). The subunit contains large cytoplasmic NH2-terminal and COOH-terminal domains. S1-S4 contributes to the voltage sensor domain (VSD) with multiple positive charges in the S4 domain acting as the primary voltage sensor for channel opening. S5-S6 in combination with the extracellular pore loop form the K^+-selective pore domain (PD). **(B)** Cumulative probability of first cardiac events from birth through age 40 years for subjects with mutations in 3 distinct of the KCNH2 channel. Adapted from Moss *et al.*[40]**.**

GENOMIC STABILITY IN HIPSC

As described in the previous sections, the advent of hiPSCs has initiated new lines of basic and translational research, opening new possibilities to study human diseases and expand the field of stem cell therapy. However, the generation of hiPSCs continues to encounter challenges in the context of genomic stability. During the different steps in generating and maintaining hiPSCs, genetic changes that range from chromosomal aberrations, copy number variations (CNVs) to point mutations, can occur[34]. For example, it has been reported that deletions of tumour-suppressor genes tend to occur with reprogramming, whereas duplications of tumour-promoting genes accumulate with passaging in culture after lines have been generated[35].

Recurrent CNVs are well documented, and several hotspots for duplication, deletions have been identified in hiPSCs (Figure 4)[36,37]. In a study with 32 hiPSC lines, researchers found that >25% of hiPSC samples have either CNVs at 1q31.3 or 17q21.1, while deletion of 8q24.3 is seen in ~12% of samples. Duplications of 20q11.21 (18%) and 2p11.2 (>25%) are also often detected in hiPSCs and are recurrently acquired in human ESCs[38,39]. A more recent meta-analysis of all hPSC genetic abnormalities reported in more than 100 publications has identified 738 recurrent genetic abnormalities mainly localized in known hotspots. The most common included trisomy 12 or 12p amplification, 20q11.21 amplification, trisomy 17 or 17q amplification, chromosome 1 amplification, and trisomy X[36]. *De novo* CNVs are also often present and are responsible for causing genetic mosaicism; however, following selective pressure they can gradually disappear during cell passaging[40].

This genomic instability raises safety concerns and could prevent the advance of hiPSC-based therapies since it is difficult to distinguish driver mutations, which leads to a positive selection of cancer cells, from passenger mutations, with no effect[41]. In

basic research, the impact of these variations is largely unknown, but they could also represent a potential barrier for accurate disease modelling with the disease phenotype in the differentiated hiPSCs being either masked or even enhanced; however, there is no consensus whether these genetic aberrations are an actual risk factor for disrupting phenotypes in culture[42].

To circumvent this problem for clinical applications, it is crucial to gain deeper insights into the genomic aberrations that occur naturally during reprogramming and prolonged culture, how selective survival may promote overgrowth of the culture of variants with certain types of mutation and to understand whether mutations can lead to truly unfavourable outcomes such as malignant outgrowth. For disease modelling, it becomes even more essential to ensure the right controls are used and the right questions asked. It is highly likely that a better understanding of genotype-phenotype relationships would have a positive impact on the robustness of the field. However, it is also plausible that some aberrations would have none or diminished effect in differentiated cells and moreover, the use of isogenic lines could circumvent these issues, at least to some extent.

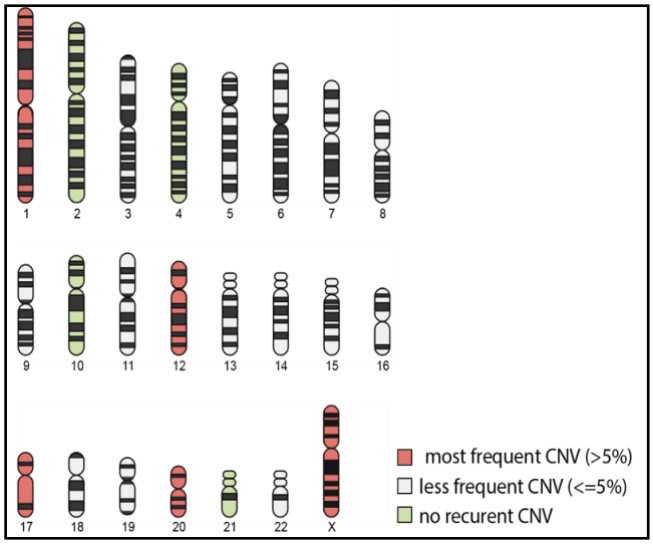

Figure 4: Ideogram summarizing the copy number variation in hiPSCs. Regions known as hot spots for genetic abnormalities represent 60% of all reported CNVs (red). Those regions are present in 6-22% of cell lines. 40% of CNVs are spread through the other 14 chromosomes in the human karyotype (white) and occur in less than 6% of cell lines. In green, no recurrent abnormalities. Figure based on Assou *et al*, 2020[32].

SCOPE OF THIS THESIS

This thesis starts with a review on the current state of the field regarding the generation of cardiomyocytes from hPSCs and methods to assess them functionally. We critically evaluated whether treatments suggested by these *in vitro* models could be translated to clinical practice, considered current shortcomings of these models, and propose methods by which they could be further improved **(chapter 2)**. We then addressed a number of these issues in subsequent experimental chapters. Firstly, we needed to establish a robust method for generating and characterising genetically modified hiPSCs–which could be used for the creation of LQT2 models **(chapter 3).** Using CRISPR/Cas9 technology, several isogenic lines harbouring mutations in *KCNH2* were generated and fully characterised **(chapter 4-6)**. As a proof-of-concept study, hiPSC-CMs with mutations either within the pore or tail region of the ion channel (Figure 3) were analysed to investigate if differences in severity observed in patients could be replicated in vitro. Electrophysiological studies confirmed differences between the mutations, with pore mutations having a more prolonged repolarisation phase and greater susceptibility to arrhythmic events when triggered by a hERG channel blocker **(chapter 4)**.

The models created from this work were also important to study the effect of compound mutations through *KCNH2* gene. Additional pairs of isogenic hiPSC lines having either the above pore or tail mutations, plus a common variant hypothesised to influence disease severity, were created. Electrophysiological studies of those lines showed that the common polymorphism KCNH2-K897T alters the phenotype of LQT2-causing *KCNH2* mutations depending on whether it is present in *cis* or *trans* with the primary mutation **(chapter 5)**. Finally, to investigate the effects of these pore and tail mutations in the absence of wildtype KCNH2, we created a pair of isogenic lines in which one allele had the missense mutation and the second allele

contained a deletion resulting in lack of protein expression from this allele. We concentrated on RNA and protein expression in these lines, investigating whether some sort of compensation mechanism was present for the difference in disease severity observed between the lines **(chapter 6).**

While generating these cell lines, a copy number variation on 1q30.3 that was not detected by standard g-band or qPCR karyotyping or karyotyping was uncovered when using a higher resolution test. The CNV was consistent in the control line at passages greater than 20, but not in all the isogenic clones derived from it. We focused on the characterisation of this duplication due to the inclusion of a key cardiomyocyte marker and its possible contribution to changes in the functional phenotype when the hiPSC-CMs were evaluated in monolayer (2D) or microtissue (3D). We also investigated the possible mechanism by which this duplication confers a growth advantage in the undifferentiated hiPSCs **(chapter 7)**.

Finally, in a general discussion of the data **(chapter 8),** I consider the importance of the findings in the thesis for the field and what hurdles still need to be overcome, not least to encourage regulatory authorities and the pharmaceutical industry to use many of these emerging resources based on human stem cells.

REFERENCES

1. Timmis, A. *et al.* European Society of Cardiology: Cardiovascular Disease Statistics 2017. *Eur. Heart J.* **39**, 508–579 (2018).
2. Takahashi, K. *et al.* Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors. *Cell* **131**, 861–872 (2007).
3. Park, I.-H. *et al.* Reprogramming of human somatic cells to pluripotency with defined factors. *Nature* **451**, 141–146 (2008).
4. Nakagawa, M. *et al.* Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. *Nat. Biotechnol.* **26**, 101–106 (2008).
5. Madrid, M., Sumen, C., Aivio, S. & Saklayen, N. Autologous Induced Pluripotent Stem Cell–Based Cell Therapies: Promise, Progress, and Challenges. *Curr. Protoc.* **1**, (2021).
6. Raab, S., Klingenstein, M., Liebau, S. & Linta, L. A Comparative View on Human Somatic Cell Sources for iPSC Generation. *Stem Cells Int.* **2014**, 1–12 (2014).
7. Bellin, M., Marchetto, M. C., Gage, F. H. & Mummery, C. L. Induced pluripotent stem cells: the new patient? *Nat. Rev. Mol. Cell Biol.* **13**, (2012).
8. Brandão, K. O., Tabel, V. A., Atsma, D. E., Mummery, C. L. & Davis, R. P. Human pluripotent stem cell models of cardiac disease: from mechanisms to therapies. *Dis. Model. Mech.* **10**, 1039–1059 (2017).
9. Savoji, H. *et al.* Cardiovascular disease models: A game changing paradigm in drug discovery and screening. *Biomaterials* **198**, 3–26 (2019).
10. Musunuru, K. Genome editing of human pluripotent stem cells to generate human cellular disease models. *Dis. Model. Mech.* (2013) doi:10.1242/dmm.012054.
11. van den Brink, L., Grandela, C., Mummery, C. L. & Davis, R. P. Concise review: Inherited cardiac diseases, pluripotent stem cells, and genome editing combined-the past, present, and future. *Stem Cells* **38**, 174–186 (2020).
12. Dugger, S. A., Platt, A. & Goldstein, D. B. Drug development in the era of precision medicine. *Nat. Rev. Drug Discov.* **17**, 183–196 (2018).
13. Carvajal-Vergara, X. *et al.* Patient-specific induced pluripotent stem-cell-derived models of LEOPARD syndrome. *Nature* **465**, 808–812 (2010).
14. Li, J., Feng, X. & Wei, X. Modeling hypertrophic cardiomyopathy with human cardiomyocytes derived from induced pluripotent stem cells. *Stem Cell Res. Ther.* **13**, 232 (2022).
15. Moretti, A. *et al.* Patient-Specific Induced Pluripotent Stem-Cell Models for Long-QT Syndrome. *N. Engl. J. Med.* **36315363**, 1397–1409 (2010).
16. van den Berg, C. W., Elliott, D. A., Braam, S. R., Mummery, C. L. & Davis, R. P. Differentiation of Human Pluripotent Stem Cells to Cardiomyocytes Under Defined Conditions. *Methods Mol. Biol.* **1353**, 163–180 (2014).
17. Kim, C. *et al.* Studying arrhythmogenic right ventricular dysplasia with patient-specific iPSCs. *Nature* **494**, 105–10 (2013).
18. Drawnel, F. M. *et al. Disease Modeling and Phenotypic Drug Screening for Diabetic Cardiomyopathy using Human Induced Pluripotent Stem Cells*. *Cell Reports* vol. 9 (2014).
19. Birket, M. J. *et al.* Contractile Defect Caused by Mutation in MYBPC3 Revealed under Conditions Optimized for Human PSC-Cardiomyocyte Function. *Cell Rep.* **13**, 733–745 (2015).
20. Giacomelli, E. *et al.* Human-iPSC-Derived Cardiac Stromal Cells Enhance Maturation in 3D Cardiac Microtissues and Reveal Non-cardiomyocyte Contributions to Heart Disease. *Cell Stem Cell* **26**, 862-879.e11 (2020).
21. Sala, L., Gnecchi, M. & Schwartz, P. J. Long QT Syndrome Modelling with Cardiomyocytes Derived from Human-induced Pluripotent Stem Cells. *Arrhythmia Electrophysiol. Rev.* **8**, 105–

110 (2019).
22. Schwartz, P. J. *et al.* Prevalence of the Congenital Long-QT Syndrome. *Circulation* **120**, 1761–1767 (2009).
23. Schwartz, P. J. *et al.* Inherited cardiac arrhythmias. *Nat. Rev. Dis. Prim.* **6**, 58 (2020).
24. Giudicessi, J. R. & Ackerman, M. J. Determinants of incomplete penetrance and variable expressivity in heritable cardiac arrhythmia syndromes. *Transl. Res.* **161**, 1–14 (2013).
25. Taggart, N. W., Haglund;, C. M., Tester, D. J. & Ackerman, M. J. Diagnostic Miscues in Congenital Long-QT Syndrome. *Circulation* **115**, 2613–2620 (2007).
26. Curran, M. E. *et al.* A molecular basis for cardiac arrhythmia: HERG mutations cause long QT syndrome. *Cell* **80**, 795–803 (1995).
27. Vandenberg, J. I. *et al.* hERG K + Channels: Structure, Function, and Clinical Significance. *Physiol. Rev.* **92**, 1393–1478 (2012).
28. Migdalovich, D. *et al.* Mutation and gender-specific risk in type 2 long QT syndrome: Implications for risk stratification for life-threatening cardiac events in patients with long QT syndrome. *Hear. Rhythm* **8**, 1537–1543 (2011).
29. Shimizu, W. *et al.* Genotype-Phenotype Aspects of Type 2 Long QT Syndrome. *J. Am. Coll. Cardiol.* **54**, 2052–2062 (2009).
30. Schwartz, P. J., Crotti, L. & George, A. L. Modifier genes for sudden cardiac death. *Eur. Heart J.* **39**, 3925–3931 (2018).
31. Sala, L., Bellin, M. & Mummery, C. L. Integrating cardiomyocytes from human pluripotent stem cells in safety pharmacology: has the time come? *Br. J. Pharmacol.* **174**, 3749–3765 (2017).
32. Chai, S. *et al.* Physiological genomics identifies genetic modifiers of long QT syndrome type 2 severity. *J. Clin. Invest.* **128**, 1043–1056 (2018).
33. Moss, A. J. *et al.* Increased Risk of Arrhythmic Events in Long-QT Syndrome With Mutations in the Pore Region of the Human Ether-a-go-go–Related Gene Potassium Channel. *Circulation* **105**, 794–799 (2002).
34. Yoshihara, M., Hayashizaki, Y. & Murakawa, Y. Genomic Instability of iPSCs: Challenges Towards Their Clinical Applications. *Stem Cell Rev. Reports* **13**, 7–16 (2017).
35. Laurent, L. C. *et al.* Dynamic Changes in the Copy Number of Pluripotency and Cell Proliferation Genes in Human ESCs and iPSCs during Reprogramming and Time in Culture. *Cell Stem Cell* **8**, 106–118 (2011).
36. Assou, S. *et al.* Recurrent Genetic Abnormalities in Human Pluripotent Stem Cells: Definition and Routine Detection in Culture Supernatant by Targeted Droplet Digital PCR. *Stem Cell Reports* **14**, 1–8 (2020).
37. Baker, D. *et al.* Detecting Genetic Mosaicism in Cultures of Human Pluripotent Stem Cells. *Stem Cell Reports* **7**, 998–1012 (2016).
38. Martins-Taylor, K. *et al.* Recurrent copy number variations in human induced pluripotent stem cells. *Nat. Biotechnol.* **29**, 488–491 (2011).
39. Spits, C. *et al.* Recurrent chromosomal abnormalities in human embryonic stem cells. *Nat. Biotechnol.* **26**, 1361–1363 (2008).
40. Hussein, S. M. *et al.* Copy number variation and selection during reprogramming to pluripotency. *Nature* **471**, 58–62 (2011).
41. Computational approaches to identify functional genetic variants in cancer genomes. *Nat. Methods* **10**, 723–729 (2013).
42. Yoshihara, M., Oguchi, A. & Murakawa, Y. Genomic Instability of iPSCs and Challenges in Their Clinical Applications. in 23–47 (2019). doi:10.1007/978-3-030-31206-0_2.